



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/595,829	03/02/2007	Jochen Peters	DE030398	7369
24737 7590 07/16/2010 PHILIPS INTELLECTUAL PROPERTY & STANDARDS P.O. BOX 3001 BRIARCLIFF MANOR, NY 10510				
EXAMINER SPOONER, LAMONT M				
ART UNIT		PAPER NUMBER		
2626				
MAIL DATE		DELIVERY MODE		
07/16/2010		PAPER		

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

### Office Action Summary

**Application No.**

10/595,829

**Applicant(s)**

PETERS, JOCHEN

**Examiner**

LAMONT M. SPOONER

**Art Unit**

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 15 May 2006.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-16 is/are rejected.
- 7) ☒ Claim(s) 1-20 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 15 May 2006 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some \* c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/GS/8)
- \_\_\_\_\_ Paper No(s)/Mail Date \_\_\_\_\_

- 4) ☐ Interview Summary (PTO-413)
- \_\_\_\_\_ Paper No(s)/Mail Date \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: \_\_\_\_\_

## **DETAILED ACTION**

### ***Introduction***

1. This office action is in response to applicant's claims filed 5/15/06. Claims 1-20 are currently pending and have been examined. There is no IDS filed. The claim to foreign priority has been acknowledged.

### ***Specification***

2. This application does not contain an abstract of the disclosure as required by 37 CFR 1.72(b). An abstract on a separate sheet is required.
3. The following guidelines illustrate the preferred layout for the specification of a utility application. These guidelines are suggested for the applicant's use.

#### **Arrangement of the Specification**

As provided in 37 CFR 1.77(b), the specification of a utility application should include the following sections in order. Each of the lettered items should appear in upper case, without underlining or bold type, as a section heading. If no text follows the section heading, the phrase "Not Applicable" should follow the section heading:

- (a) TITLE OF THE INVENTION.
- (b) CROSS-REFERENCE TO RELATED APPLICATIONS.
- (c) STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT.
- (d) THE NAMES OF THE PARTIES TO A JOINT RESEARCH AGREEMENT.
- (e) INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED ON A COMPACT DISC.
- (f) BACKGROUND OF THE INVENTION.
  - (1) Field of the Invention.

- (2) Description of Related Art including information disclosed under 37 CFR 1.97 and 1.98.
- (g) BRIEF SUMMARY OF THE INVENTION.
- (h) BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S).
- (i) DETAILED DESCRIPTION OF THE INVENTION.
- (j) CLAIM OR CLAIMS (commencing on a separate sheet).
- (k) ABSTRACT OF THE DISCLOSURE (commencing on a separate sheet).
- (l) SEQUENCE LISTING (See MPEP § 2424 and 37 CFR 1.821-1.825. A "Sequence Listing" is required on paper if the application discloses a nucleotide or amino acid sequence as defined in 37 CFR 1.821(a) and if the required "Sequence Listing" is not submitted as an electronic document on compact disc).

#### **Content of Specification**

- (a) Title of the Invention: See 37 CFR 1.72(a) and MPEP § 606. The title of the invention should be placed at the top of the first page of the specification unless the title is provided in an application data sheet. The title of the invention should be brief but technically accurate and descriptive, preferably from two to seven words may not contain more than 500 characters.
- (b) Cross-References to Related Applications: See 37 CFR 1.78 and MPEP § 201.11.
- (c) Statement Regarding Federally Sponsored Research and Development: See MPEP § 310.
- (d) The Names Of The Parties To A Joint Research Agreement: See 37 CFR 1.71(g).
- (e) Incorporation-By-Reference Of Material Submitted On a Compact Disc: The specification is required to include an incorporation-by-reference of electronic documents that are to become part of the permanent United States Patent and

Trademark Office records in the file of a patent application. See 37 CFR 1.52(e) and MPEP § 608.05. Computer program listings (37 CFR 1.96(c)), "Sequence Listings" (37 CFR 1.821(c)), and tables having more than 50 pages of text were permitted as electronic documents on compact discs beginning on September 8, 2000.

- (f) Background of the Invention: See MPEP § 608.01(c). The specification should set forth the Background of the Invention in two parts:
- (1) Field of the Invention: A statement of the field of art to which the invention pertains. This statement may include a paraphrasing of the applicable U.S. patent classification definitions of the subject matter of the claimed invention. This item may also be titled "Technical Field."
  - (2) Description of the Related Art including information disclosed under 37 CFR 1.97 and 37 CFR 1.98: A description of the related art known to the applicant and including, if applicable, references to specific related art and problems involved in the prior art which are solved by the applicant's invention. This item may also be titled "Background Art."
- (g) Brief Summary of the Invention: See MPEP § 608.01(d). A brief summary or general statement of the invention as set forth in 37 CFR 1.73. The summary is separate and distinct from the abstract and is directed toward the invention rather than the disclosure as a whole. The summary may point out the advantages of the invention or how it solves problems previously existent in the prior art (and preferably indicated in the Background of the Invention). In chemical cases it should point out in general terms the utility of the invention. If possible, the nature and gist of the invention or the inventive concept should be set forth. Objects of the invention should be treated briefly and only to the extent that they contribute to an understanding of the invention.

- (h) Brief Description of the Several Views of the Drawing(s): See MPEP § 608.01(f). A reference to and brief description of the drawing(s) as set forth in 37 CFR 1.74.
- (i) Detailed Description of the Invention: See MPEP § 608.01(g). A description of the preferred embodiment(s) of the invention as required in 37 CFR 1.71. The description should be as short and specific as is necessary to describe the invention adequately and accurately. Where elements or groups of elements, compounds, and processes, which are conventional and generally widely known in the field of the invention described and their exact nature or type is not necessary for an understanding and use of the invention by a person skilled in the art, they should not be described in detail. However, where particularly complicated subject matter is involved or where the elements, compounds, or processes may not be commonly or widely known in the field, the specification should refer to another patent or readily available publication which adequately describes the subject matter.
- (j) Claim or Claims: See 37 CFR 1.75 and MPEP § 608.01(m). The claim or claims must commence on separate sheet or electronic page (37 CFR 1.52(b)(3)). Where a claim sets forth a plurality of elements or steps, each element or step of the claim should be separated by a line indentation. There may be plural indentations to further segregate subcombinations or related steps. See 37 CFR 1.75 and MPEP § 608.01(i)-(p).
- (k) Abstract of the Disclosure: See MPEP § 608.01(f). A brief narrative of the disclosure as a whole in a single paragraph of 150 words or less commencing on a separate sheet following the claims. In an international application which has entered the national stage (37 CFR 1.491(b)), the applicant need not submit an abstract commencing on a separate sheet if an abstract was published with the international application under PCT Article 21. The abstract that appears on the cover page of the pamphlet published by the International Bureau (IB) of the

World Intellectual Property Organization (WIPO) is the abstract that will be used by the USPTO. See MPEP § 1893.03(e).

- (l) Sequence Listing. See 37 CFR 1.821-1.825 and MPEP §§ 2421-2431. The requirement for a sequence listing applies to all sequences disclosed in a given application, whether the sequences are claimed or not. See MPEP § 2421.02.

### ***Claim Objections***

4. Claims 1, 3, 8, 9, 10, 11, 13, 17, 18, and 20 are objected to because of the following informalities:

In claim 1, lines 2 and 3, applicant claims "featuring a plurality of text units (320, 322, ...)". The references to the drawings contain ellipsis which imply a multiple drawing elements, however Fig. 3 contains many drawing elements (i.e. 302-366) and they are all not text units. The Examiner notes, the references to the drawings are properly enclosed in parentheses and do not affect the scope of the claims as per MPEP (608.01(m)), see below, however, the ellipses within the references to the drawings require clarification.

The MPEP (608.01 (m)) states:

**"Reference characters corresponding to elements recited in the detailed description and the drawings may be used in conjunction with the recitation of the same element or group**

**of elements in the claims. The reference characters, however, should be enclosed within parentheses so as to avoid confusion with other numbers or characters which may appear in the claims. The use of reference characters is to be considered as having no effect on the scope of the claims."**

The Examiner notes the ellipsis are found in at least one location in claims 1, 3, 8, 9, 10, 11, 13, 17, 18, and 20.

The respective dependent claims are objected to as being dependent on an objected independent parent claim.

Appropriate correction is required.

***Claim Rejections - 35 USC § 101***

5. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

6. Claim 1-10 are rejected under 35 USC 101 as not falling within one of the four statutory categories of invention. While the claims recite a series of steps or acts to be performed, regarding claim 1,

"A method of text clustering for the generation of language models, a text (300) featuring a plurality of text units (320, 322 .... ), each of which



having at least one word (302, 304,...), the method of text clustering comprising the steps of:

- assigning each of the text units (320, 322,...) to one of a plurality of provided clusters (330, 332 .... ),
- determining for each text unit a set of emission probabilities (340, 350), each emission probability (342, 344,...,352, 354,...) being indicative of a correlation between the text unit (320, 322 .... ) and a cluster (330, 332,...), the set of emission probabilities being indicative of the correlations between the text unit and the plurality of clusters,
- determining a transition probability (362, 364,...) being indicative that a first cluster (330) being assigned to a first text unit (320) in the text is followed by a second cluster (332) being assigned to a second text unit (322) in the text, the second text unit (322) subsequently following the first text unit (320) within the text,
- performing an optimization procedure based on the emission probability and the transition probability in order to assign each text unit to a cluster."

a statutory "process" under 35 USC 101 must (1) be tied to another statutory category (such as a manufacture or a machine), or (2) transform

underlying subject matter (such as an article or material) to a different state or thing. The instant claim(s) neither transform underlying subject matter nor positively recite structure associated with another statutory category, and therefore do not define a statutory process (i.e. a limitation in the body of the claim citing, implemented by computer, processor, loaded from memory, etc., as supported by applicants disclosure, would tie the method to another statutory category).

Claims 2-10 do not rectify the above rejection and are thus rejected under the same rationale.

7. The claimed invention, regarding claims 11-16, is directed to non-statutory subject matter. More specifically, claim 11 merely cites functional descriptive material, a computer program per se, without any embodiment. More specifically, claim 11 cites,

"A computer program product for text clustering for the generation of language models, a text (300) featuring a plurality of text units (320, 322,...), each of which having at least one word (302, 304, ...), the computer program product comprising program means for:

"assigning each of the text units (320, 322,...) to one of a plurality of provided clusters (330, 332,...),

- determining for each text unit a set of emission probabilities (340, 350), each emission probability (342, 344,..., 352, 354,...) being indicative of a correlation between the text unit (320, 322,...) and a cluster (330, 332,...), the set of emission probabilities being indicative of the correlations between the text unit and the plurality of clusters,

- determining a transition probability (362, 364,...) being indicative that a first cluster (330) being assigned to a first text unit (320) in the text is followed by a second cluster (332) being assigned to a second text unit (322) in the text, the second text unit (322) subsequently following the first text unit (320) within the text,

- performing an optimization procedure based on the emission probability and the transition probability in order to assign each text unit to a cluster.”

The Examiner notes the “product” in the broadest reasonable interpretation appears to be a computer program, wherein the applicant provides no description in the disclosure of what the “computer program product” is. Therefore, the descriptions or expressions of the program are not deemed physical “things.” They are neither computer components nor statutory processes, as they are not “acts” being performed. Such claimed

computer programs do not define any structural and functional interrelationships between the computer program and other claimed elements of a computer which permit the computer program's functionality to be realized. In contrast, for example, a claimed computer-readable medium encoded with a computer program is a computer element which defines structural and functional interrelationships between the computer program and the rest of the computer which permit the computer program's functionality to be realized, and is thus statutory. See Lowry, 32 F.3d at 1583-84, 32 USPQ2d at 1035.

Claims 12-16 do not rectify the current rejection and are each also rejected under the same rationale.

***Claim Rejections - 35 USC § 112***

8. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

9. Claim 11-16 are rejected under 35 U.S.C. 112, first paragraph, because the specification, while being enabling for at most only the means as disclosed by the inventor (see specification p.13 lines 11-23), does not reasonably provide enablement for a single means which covers every

conceivable means for achieving applicant's claims (claims 11-16). The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make the invention commensurate in scope with these claims. Wherein the means recitation does not appear in combination with another recited element of means, and is subject to undue breadth (see MPEP 2164.08 (a)).

***Allowable Subject Matter***

10. Claims 1-10 would be allowable if rewritten or amended to overcome the rejection(s) under 35 USC 101, claims 11-16 would be allowable if rewritten to overcome the rejections under 35 USC 101 and 35 U.S.C. 112, 1<sup>st</sup> paragraph, as set forth in this Office action. Claims 17-20 would be allowable if claims 17, 18 and 20 were rewritten to overcome the respective objections.

11. The following is a statement of reasons for the indication of allowable subject matter:

The instant application is deemed to be directed to a non-obvious improvement over the invention patented Moreno et al. (Moreno, US 6,772,120), Dom et al. (Dom, US 6,584,456), Conklin (US 6,415,283), and Bangalore et al. (Bangalore, US 6,415,248). Moreno teaches determining

observation emission probabilities, training segment clusters to find transition probabilities, building smoothed language models. Dom teaches optimizing clustering of documents/text units based on cluster counts and subsets of clusters. Conklin teaches optimizing clustering techniques to identify clusters. Bangalore teaches an optimizing clustering in order to build complex linguistic models from a corpus.

Neither Moreno, Dom, Conklin or Bangalore alone or in obvious combination teach:

Regarding claims 1 and 11:

“- determining for **each text unit** a **set** of emission probabilities (340, 350), each emission probability (342, 344, ..., 352, 354, ...) being indicative of a correlation between the text unit (320, 322,...) and a cluster (330, 332,...), **the set of emission probabilities** being indicative of the correlations between **the text unit** and **the plurality of clusters**,

- determining a transition probability (362, 364,...) being indicative that a **first cluster** (330) **being assigned to a first text unit** (320) **in the text is followed by a second cluster** (332) **being assigned to a second text unit** (322) **in the text, the second text unit** (322) **subsequently following the first text unit** (320) **within the text**,

- performing an **optimization procedure based on the emission probability and the transition probability** in order to assign **each text unit to a cluster.**"

Regarding claim 17, "- means for determining for **each text unit a set** of emission probabilities (340, 350), each emission probability (342, 344, ..., 352, 354, ...) being indicative of a correlation between the text unit (320, 322,...) and a cluster (330, 332,...), **the set of emission probabilities** being indicative of the correlations between **the text unit** and **the plurality of clusters**,

- means for determining a transition probability (362, 364,...) being indicative that a **first cluster (330) being assigned to a first text unit (320) in the text is followed by a second cluster (332) being assigned to a second text unit (322) in the text, the second text unit (322) subsequently following the first text unit (320) within the text,**

-means for performing an **optimization procedure based on the emission probability and the transition probability** in order to assign **each text unit to a cluster.**"

Dependent claims 2-10, 12-16 and 18-20 would be allowed as they inherit the allowable subject matter of their respective independent parent claim.

### ***Conclusion***

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- Gao et al. (US 7,275,029) teaches clustering and optimization of language model.
- Ushioda (US 5,835,892) teaches class based clustering, wherein adjacent words of a first class/cluster and second class/cluster, are utilized to optimize word clustering result.
- Keung et al. (us 2002/0193981) teaches incremental and interactive clustering on high-dimensional data, achieving optimal clustering by dimension reduction.
- Mishara et al. (US 7,739,313) teaches finding conjunctive clusters.
- Gaussier et al. (US 7,644,102) teaches emission probability determination, clustering by an annealing Expectation Maximization algorithm.



- Bharat et al. (US 7,568,148) teaches clustering data content y topic, sorting, ranking and optimizing the clusters.
- Marchisio (US 6,510,406) teaches generating structured content by clustering and optimizing the clusters in a vector space.
- Burdick et al. (US 7,185,001) teaches iterative clustering and reclustering documents into categories, thus optimizing the clustering procedure.
- Vaithyanathan et al. (US 5,857,179) teaches re-clustering text units, thus optimizing the cluster.
- Blei et al., Topic Segmentation with an Aspect Hidden Markov Model, teaches combining emission probabilities with transition probabilities, thus accounting for text cohesion in clustering.
- Brants et al., Topic-based document segmentation with probabilistic latent semantic analysis, teaches clustering text units and an optimization procedure for optimal cluster determination.
- Woszcyna et al. Inferring Linguistic Structure in Spoken language, teaches determining both emission and transition

probabilities in developing a language model based on  
cluster/topic based text units.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to LAMONT M. SPOONER whose telephone number is (571)272-7613. The examiner can normally be reached on 8:00 AM - 5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571/272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Lamont M Spooner/  
Examiner, Art Unit 2626

7/9/10